

# Inhoudsopgave

Week 1:	College Werkgroep
Week 2:	College Werkgroep
Week 3:	College Werkgroep
Week 4:	College Werkgroep
Week 5:	College Werkgroep
Week 6:	College Werkgroep
Week 7:	College Werkgroep

# Week 1: College: Correlaties en maten voor effectgrootte

## Correlatie en Causaliteit

Correlatie is de samenhang tussen variabelen. Twee variabelen correleren met elkaar als het ene tegelijkertijd optreedt met het andere. Het zegt niet meteen dat er ook een oorzaak-gevolgrelatie is tussen 2 variabelen: het zegt dat er een bepaalde samenhang is.

De correlatie ligt tussen de -1 en de 1. Wanneer de correlatie 0 is, is er geen samenhang tussen variabelen.

Causaliteit gaat over oorzaak-gevolgrelaties. Het is de sleutel tot het begrijpen van de werkelijkheid.

Er zijn 3 criteria voor causaliteit:

1. Samenhang
2. Opeenvolging in tijd van oorzaak en gevolg
3. Uitsluiting van alternatieve verklaringen

Het uitsluiten van alternatieve verklaringen is erg lastig. Daarom is het moeilijk om aan alle criteria te voldoen.

Causaliteit vaststellen gaat het beste in experimenteel onderzoek, en enigszins in correlatieve onderzoek.

## Pearson (product-moment) correlatie r:

De Pearson correlatie is van interval meetniveau. Hierbij wordt gerekend met gemiddelden. Met deze correlatie kun je uitdrukken hoe sterk de lineaire relatie tussen 2 variabelen is. Dit doe je aan de hand van de volgende formule:

$$r_{xy} = r = \frac{1}{n-1} \sum \left( \underbrace{\frac{x_i - \bar{x}}{s_x}}_{z_x} \right) \left( \underbrace{\frac{y_i - \bar{y}}{s_y}}_{z_y} \right) = \frac{1}{n-1} \sum z_x z_y$$

Oftwel:  $r_{xy} = (\sum z_x z_y) / (n-1)$

De sterkte van de samenhang hier wordt uitgedrukt in een getal dat onafhankelijk is van lokatie en schaal (meeteenheid) van de twee variabelen x en y.

Voor  $r_{xy}$  geldt:

1. Samenhang ( $r$ ) ligt altijd tussen -1 en 1
2. Als er geen *lineaire* associatie is, dan is  $r$  gelijk aan 0
3.  $r$  meet alleen de mate van *lineair* verband, ook al suggereert de scatterplot een ander, *niet lineair* verband
4.  $r$  is niet robuust tegen uitbijters

### De Spearman rangorde correlatie:

Deze wordt ook wel uitgedrukt in  $\rho$  (rho) of  $r_s$ . Het gaat hierbij om het lineaire verband tussen 2 variabelen van ordinaal meetniveau.

De Spearman rangorde correlatie heeft de volgende eigenschappen:

- a. Het controleren van de overeenkomst tussen 2 testen (interrater reliability)
- b. Het benadert variabelen waarbij de gemeten precieze score buiten beschouwing moet worden gelaten.
- c. Het is een robuuste variant van Pearson  $r$ .

De formule voor Spearman's rho is:

$$r_s = 1 - (6 \cdot \sum D^2) / (n^3 - n)$$

Als alle rangorde's identiek is, zijn alle  $D$ 's nul, en is de Spearman correlatie dus 1.

### De Punt-biseriële correlatie:

Deze wordt ook wel uitgedrukt in  $r_{pb}$ . Het is de relatie tussen een binaire variabele ( $X$ ) (een variabele met maar 2 uitkomstmogelijkheden) en een variabele op interval meetniveau ( $Y$ ).

De formule voor deze correlatie is:

$$r_{pb} = ((y(\text{gem}1) - y(\text{gem})) / S_y) \cdot S_x$$

### De phi-coëfficiënt $\Phi$ :

Hierbij wordt de relatie tussen 2 binaire variabelen weergegeven. Je doet dit door middel van een kruistabel. In die kruistabel zijn 4 combinaties mogelijk (2rijen\*2kolommen). Deze geef je de letters A,B,C,D. Daarna is als volgt de phi-coëfficiënt te berekenen:

$$\Phi = ((A \cdot D) - (B \cdot C)) / \sqrt{((A+B)(C+D)(A+C)(B+D))}$$

De phi-coëfficiënt heeft een relatie met de  $X^2$  :

$$\Phi = \sqrt{X^2 / n}$$

Een overzichtje van verschillende variabelen en de bijpassende correlaties:

Bij twee kwantitatieve variabelen gebruik je de Pearson  $r$ . Bij twee ordinale variabelen gebruik je Spearman's  $\rho$ . Bij twee dichotome variabelen gebruik je de Phi-coëfficiënt. En bij een dichotome en een kwantitatieve variabele gebruik je de punt-biseriële correlatie  $r_{pb}$ .

**De t-toets :**

De statistische significantie van de samenhang wordt bepaald met een t-toets.

Deze toetst de nulhypothese  $H_0: \rho = 0$  en bijvoorbeeld  $H_a: \rho > 0$

Je berekent de t-toets met:  $t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}}$  met  $df = n-2$ .

In de t-tabel kun je bij de gewenste  $\alpha$  en het aantal vrijheidsgraden de  $t^*$  opzoeken. Hoe kleiner de p-waarde, hoe groter het bewijs tegen  $H_0$  is.

Als de correlatie groter is, dan wordt ook de t-waarde groter. Dit kun je aan de formule zien, want  $r$  staat boven de breukstreep.

**Effectgrootte**

De waarde van de correlatie geeft de sterkte van de samenhang aan maar die waarde is moeilijk te interpreteren omdat je niet weet wat bijvoorbeeld  $r = .60$  betekent. Je kunt het ook moeilijk vergelijken,  $r = .60$  is niet een twee keer zo sterke samenhang als  $r = .30$ . Daarom kun je  $r$  kwadrateren, dit heet de coëfficiënt of determination (COD) of Proportion of Variance Explained (VAF). Hiermee wordt de richting niet meer zichtbaar, doordat je met het kwadrateren automatisch een positief getal vindt.

Naast een maat voor sterkte van samenhang is dus ook een maat voor *effectgrootte* nodig. Effect heeft altijd te maken met samenhang. Bij een causale relatie is het ene verschijnsel oorzaak en het andere het gevolg. Men kan ook zeggen dat het ene verschijnsel invloed (of effect) heeft op het andere. Hiervoor gebruiken we twee alternatieve maten om effectgrootte te berekenen:

$$1. \text{ Cohen's } d \quad \square \text{ op basis van populatiewaarden} \quad = \frac{\mu_1 - \mu_2}{\sigma}$$

$$2. \text{ Hedges' } g \quad \square \text{ op basis van steekproefsituatie} \quad = \frac{\bar{X}_1 - \bar{X}_2}{s_p}$$

In de tabellen op sheet 34 worden de vuistregels gegeven voor de sterkte van relatie en grootte van effect. Deze kun je gebruiken als algemene richtlijnen.

## Week 1: Werkgroep

### Opdracht 1.1:

- We verwachten een positief verband
- De Pearson correlatie  $r$  berekenen we hier met behulp van de formule  $r_{xy} = (\sum z_x z_y) / (n-1)$ . In de tabel in het werkboek kan je de z-scores vinden, door die te vermenigvuldigen en daarna op te tellen komen we op een  $\sum z_x z_y$  van 5,5456. Als we dan de formule invullen krijgen we  $5,5456 / (8-1) = 0,79$ .
- Ho:  $\rho = 0$ . Ha:  $\rho > 0$ .
- We gebruiken een t-toets om de nulhypothese te toetsen met  $t = (r\sqrt{(n-2)}) / (\sqrt{1-r^2})$  met  $df=n-2$ . Dit geeft  $t(6) = (0,79 * \sqrt{(8-2)}) / (\sqrt{1-0,79^2}) = 3,18$ . De p-waarden van deze t liggen tussen de 0,01 en de 0,005, dit is onder de alpha en dus moet Ho worden verworpen.
- We mogen niet concluderen dat er een oorzakelijk verband is.

### Opdracht 1.2:

- sdjkgghs
- Er zijn twee manieren om  $r_s$  te berekenen. De eerst manier gaat via de specifieke formule van  $r_s$ . Deze is  $r_s = 1 - (6 * \sum D^2) / (n^3 - n)$ . Als we de formule invullen krijgen we  $r_s = 1 - (6 * 12) / (8^3 - 8) = 0,8571$ .
- Bij de tweede manier worden z-scores gebruikt. Via de tabel in het werkboek is de kruisproductensom van de z-scores te berekenen. Invullen in de uniforme formule geeft  $r_s = (\sum z_x z_y) / (n-1) = 6,00 / (8-1) = 0,8571$ .
- Ho:  $\rho = 0$  en Ha:  $\rho > 0$ .
- We gebruiken een t-toets met  $t = (r\sqrt{(n-2)}) / (\sqrt{1-r^2}) = (0,8571 * \sqrt{(8-2)}) / (\sqrt{1-0,8571^2}) = 4,076$ . Als we deze t-waarde opzoeken in de t-tabel, dan zien we dat de p tussen de 0,005 en de 0,0025 ligt. Dit betekent dat we Ho kunnen verworpen.

### Opdracht 1.3:

- De specifieke formule voor de punt-biseriële correlatie is  $r_{pb} = ((y(\text{gem1}) - y(\text{gem})) / S_y) * S_x$ . Als we die invullen krijgen we  $r_{pb} = ((4,5 - 7,75) / 2,475) * 0,535 = -0,702$ .
- De tweede manier:  $r_{pb} = (\sum z_x z_y) / (n-1) = -4,9135 / (8-1) = -0,702$ .
- De  $r_{pb}$  is negatief, dit betekent dat de richting ook negatief is. Mannen hebben dan dus hogere scores dan vrouwen.
- Ho:  $\rho = 0$  en Ha:  $\rho < 0$ .
- $t = (r\sqrt{(n-2)}) / (\sqrt{1-r^2}) = (-0,702 * \sqrt{(8-2)}) / (\sqrt{1-(-0,702)^2}) = -2,41$ . De p-waarden liggen voor deze t tussen de 0,025 en 0,05, en dus wordt Ho verworpen.

## Opdracht 1.4.

	X=0	X=1	TOTAAL
Y=0	3 (A)	1 (B)	4 (A+B)
Y=1	1 (C)	3 (D)	4 (C+D)
TOTAAL	4 (A+C)	4 (B+D)	8 (N)

- De specifieke formule van de Phi-coefficient is:  $\Phi = ((A*D)-(B*C)) / \sqrt{((A+B)(C+D)(A+C)(B+D))}$ . Als we die invullen krijgen we  $\Phi = (9-1) / (\sqrt{4*4*4*4}) = 0.5$ .
- Als de de Phi-coefficient berekenen met behulp van z-scores krijgen we  $\Phi = (\sum z_x z_y) / (n-1) = 3,5 / (8-1) = 0.5$ .
- Het is een positieve waarde en dus ook een positieve relatie.
- Het gaat om een eenzijdig toetsprobleem. Ho:  $\rho = 0$  en Ha:  $\rho > 0$ .
- $t = (r\sqrt{(n-2)}) / (\sqrt{1-r^2}) = (0,5 * (8-2)) / (\sqrt{1-0,5^2}) = 1,41$ . De p-waarde voor  $t=1,41$  ligt tussen de 0,15 en de 0,10. Aangezien de alpha 0,05 is, wordt Ho behouden.
- $X^2$  bereken je met de formule  $\Phi = \sqrt{(X^2 / n)}$ . Als je deze ombouwd krijg je  $X^2 = \Phi^2 * n = 8 * 0,25 = 2$ .
- Gebruik voor het toetsen van  $X^2$  de  $X^2$ -tabel. De vrijheidsgraden zijn gelijk aan  $df = (kolom-1)(rij-1) = (2-1)(2-1) = 1$ . Ho wordt dan behouden, aangezien de p-waarden tussen de 0,20 en 0,15 liggen.

## Opdracht 1.5:

- $\Phi = 0.5$ . Dit is een large effectsize. Helaas in de n zo klein dat de power niet groot genoeg is.
- Bij gebruik van de punt-biseriele correlatie als maat voor de effectgrootte is de effectisize ook groot. Want  $r_{pb} = 0,702$  (het minteken mag hier weggelaten worden).
- Een t-toets voor onafhankelijke steekproeven ziet er hetzelfde uit als wat we hebben gedaan bij opdracht 1.3. T is dan:  $t = (r\sqrt{(n-2)}) / (\sqrt{1-r^2}) = (-0,702 * (\sqrt{8-2})) / (\sqrt{1-(-0,702)^2}) = -2,41$ .

## Week 2: College: Enkelvoudige Lineaire Regressie

### Regressiemodellen

Regressiemodellen voorspellen oorzaak en gevolgrelaties. Het zijn modellen van causale samenhang met een relatie tussen de voorspellende variabele (x) en de afhankelijke variabele (y).

### Lineaire Regressie

Regressieanalyse: Het opstellen van een regressievergelijking. Een variabele voorspellen uit een andere variabele. Dus y voorspellen uit x. Variabele Y is de responsvariabele (predicted variabele) en de x variabele is de predictor variabele.

De formule voor een regressielijn van de steekproef is als volgt:  $\hat{y} = b_0 + b_1x$ . Deze formule bevat een helling en een intercept.

- De helling van de formule is de helling of het regressiegewicht. Deze bepaalt hoe steil de lijn zal lopen. In de formule wordt de helling aangeduid met  $b_1$ .
- Het intercept is de plaats waar de lijn zal beginnen (waar deze de y-as kruist). Dit is niet altijd bij 0. In de formule wordt het intercept aangeduid met  $b_0$ .

Intercept en helling (ook wel slope genoemd) worden regressiecoëfficiënt genoemd.

Voor de populatie is de formule voor een regressielijn als volgt:  $\mu_y = \beta_0 + \beta_1x$

De regressielijn beschrijft de gemiddelde respons voor iedere subpopulatie als functie van x. Dus iedere waarde op de regressielijn is het gemiddelde van een subpopulatie. Hiermee wordt een rechte lijn geplot. Binnen iedere subpopulatie zijn de y- waarden normaal verdeeld.

Echter is er een verschil tussen de waargenomen waarden van de subpopulaties en de gemiddelden die worden weergegeven op de regressielijn. Het verschil tussen deze waargenomen waarde en de gemiddelde waarde kan men uitdrukken in het residu:  $\epsilon$ . Als de samenhang hoog is, is het residu klein.

### **Standaardfout voor $b_0$ en $b_1$ :**

Op sheet 9 zijn de formules voor deze berekeningen te zien. Hiermee kun je de standaardfout van  $b_0$  en  $b_1$  berekenen. Aan de hand van deze standaardfout is de standaarddeviatie te benaderen. Het intercept wordt haast nooit getoetst, omdat hier geen vragen voor zijn. Er wordt namelijk bijna geen onderzoek naar het intercept gedaan. De helling zegt meer over de lijn en wordt daarom vaker getoetst. Je gebruikt dus bijna altijd de formule voor  $b_1$ . Wanneer we te maken hebben met een standaardfout, gaan we automatisch denken aan de t-toets. Dit is namelijk de toets die met een benaderde standaarddeviatie werkt. Daarom kunnen we deze hierbij weer gebruiken.

De volgende rekestip is van belang om te onthouden, met de onderstaande formule kun je de kwadratensom gemakkelijker berekenen:

$$\sum(x_i - \bar{x})^2 = (n - 1)s_x^2$$

### **Betrouwbaarheidsinterval en Significantietoets**

Heb je eenmaal het regressiegewicht gevonden dan kun je het betrouwbaarheidsinterval berekenen waarbinnen dit ligt. Zo kun je met een bepaalde zekerheid zeggen tussen welke waarden dit ligt.

### **ANOVA-tabel:**

De ANOVA-tabel is de tabel waarin je de aspecten van regressieanalyse terugvindt. Hieronder kun je deze tabel vinden:



	Geobserveerd	=	Verklaard	+	
	DATA	=	FIT	+	RESIDUAL
Data	$y_i$	=	$\hat{y}_i$	+	$e_i$
Deviations	$(y_i - \bar{y})$ total variation	=	$(\hat{y}_i - \bar{y})$ variation fitted	+	$(y_i - \hat{y}_i)$ variation about fitted line  Verschil tussen geobserveerd en geschat
Sums of squares	$\sum (y_i - \bar{y})^2$	=	$\sum (\hat{y}_i - \bar{y})^2$	+	$\sum (y_i - \hat{y}_i)^2$
	SST	=	SSM	+	SSE
Voor df's analoog	DFT	=	DFM	+	DFE
	(n-1)	=	1	+	(n-2)
	SST/DFT = MST		SSM/DFM = MSM		SSE/DFE = MSE
	variantie		verklaarde variantie		Error - variantie rond regressielijn

### De F-toets en regressieanalyse:

De f-toets is ook weer te gebruiken bij regressieanalyse. Je kunt deze berekenen middels de volgende formule:  $F = \text{MSM} / \text{MSE}$ . MSM is de verklaarde variantie. Deze kan men berekenen door de SS van het model / DF van het model. De MSE is de variantie rond de regressielijn. Deze kan men berekenen door de SS van de error/ DF van de error. Met de uitkomsten kun je de F- waarde berekenen en deze hangt weer samen met de t- waarde:  $F = t^2$ .

Als je dus een F- waarde van 3 vind, kun je deze omzetten naar een t- waarde. De t- waarde die hierbij hoort zal 9 zijn.

De proportie verklaarde variantie (VAF) kan je berekenen door  $SSM/SST$ . Dit is hetzelfde als  $r_{xy}^2$ .

## Week 2: Werkgroep

### Opdracht 2.1:

- Stressbestendigheid is hier de response variabele
- De populatie is het Koninklijke Instituut voor de Marine.
- De regressielijn van de populatie ziet er zo uit:  $\mu y = \beta_0 + \beta_1 x$ . Het statistische model is:  $\mu y = \beta_0 + \beta_1 x + \varepsilon$
- $\beta_0$ ,  $\beta_1$  en  $\varepsilon$  zijn de parameters van dit model. Spreek  $\varepsilon$  uit als sigma (dit is de standaarddeviatie van de error).
- Het is een positief, vrij sterk verband.
- Het regressiegewicht  $b_1$  berekenen we met behulp van de formule:  $b_1 = r * (s_y / s_x) = 0,8856 * (\sqrt{22,5} / \sqrt{106}) = 0,408$ .  
Het intercept berekenen we met  $b_0 = \bar{y} - b_1 x_{gem} = 40,0 - 0,408 * 38,0 = 24,495$
- $\hat{Y} = 24,495 + 0,408x$

### Opdracht 2.2:

- $H_0: \beta_1 = 0$  en  $H_a: \beta_1 > 0$
- We gebruiken hier een t-toets.
- De errorterm  $e_i$  is hetzelfde als  $y - \hat{y}$ .
- De variantie van de errortermen berekenen we met behulp van de formule  $s^2 = \sum (y - \hat{y})^2 / (n-2) = 19,41 / 3 = 6,47$ . De schatting van de parameter sigma,  $s$ , is dan  $\sqrt{6,47} = 2,54$ .
- De nulhypothese toetsen we met de formule:  $t = b / SE_{b_1}$ . Hierbij is  $SE_{b_1} = s / (\sqrt{\sum (x_i - x_{gem})^2}) = s / (\sqrt{(n-1) * s_x^2}) = 2,54 / (\sqrt{4 * 106}) = 0,1235$ . De t-waarde is dan  $t = 0,408 / 0,1235 = 3,303$ . De vrijheidsgraden zijn  $n-2 = 3$ . De p-waarden kunnen we dan opzoeken in de tabel, en dan zien we dat deze liggen tussen de 0,025 en 0,02. Het is een eenzijdige toets dus  $H_0$  mag worden verworpen.
- Hier gelden dezelfde hypothesen, maar met  $t = (r \sqrt{(n-2)}) / (\sqrt{1-r^2})$ . De  $r$  is gegeven in opdracht 2.1. De formule invullen geeft  $t = (0,8856 * (\sqrt{5-2})) / (\sqrt{1-0,8856^2}) = 3,303$ .

### Opdracht 2.3:

- $x^* = 35$  in de formule invullen geeft  $\mu y = 24,495 + 0,408 * 35 = 38,78$ . Het betrouwbaarheidsinterval ziet er zo uit:  $b_1 \pm t^* SE_{b_1}$ . Hierbij haal je  $t^*$  uit de tabel. Het betrouwbaarheidsinterval is hier:  $[34,97; 42,58]$ .
- Het 95% betrouwbaarheidsinterval is hier:  $[29,83; 47,72]$ .
- Bij Jan met een score van  $x^* = 50$  wordt het 95% betrouwbaarheidsinterval groter, want je gaat verder van het gemiddelde af.

### Opdracht 2.4:

- De conclusie zou ongeveer hetzelfde zijn, maar de richtingen blijven met de ANOVA toetsing niet bewaard. Er is dan dus wel een verband, maar je weet niet in welke richting dat verband is.
- $H_0: \beta_1 = 0$  en  $H_a: \beta_1 \neq 0$
- Hier moet de F-toets worden gebruikt
- SSM berekenen we door de tabel in het werkboek in te vullen en dan de kolom met  $(\hat{y} - \bar{y})^2$  op te tellen. We krijgen dan een SSM van 70,59.
- $SST = \sum (y_i - \bar{y})^2 = (n-1) * s_y^2$ . Invullen geeft  $(5-1) * 22,5 = 90$ . SSE is dan dus  $SST - SSM = 90 - 70,59 = 19,41$ .
- F heeft een waarde van 10,91.
- Als je een F van 10,91 opzoekt in de tabel met  $df=1$  en  $df=3$ , dan krijg je een p-waarde van tussen de 0,050 en 0,025. Het is wel een tweezijdige toets maar dan nog is de p-waarde onder de alpha, dus kunnen we  $H_0$  verworpen.

- h. Verwerp  $H_0$ .
- i.  $F = t^2 = 3,303^2 = 10,91$ . De relatie tussen de F-waarde en de t-waarde klopt dus.
- j.  $R^2 = VAF = SSM/SST = 70,59 / 90 = 0,7843$
- k.  $r = \sqrt{R^2} = \sqrt{0,7843} = 0,8856$ . Dit is precies wat is gegeven in opdracht 2.1.
- l. De proportie onverklaarde variantie is  $1 - r^2 = 1 - (SSM/SST) = 1 - 0,7843 = 0,2157$ .
- m.  $SSE = (1 - r^2) * SST = 0,2157 * 90 = 19,41$ .

## Week 3: College: Meervoudige Lineaire Regressie

### Multipale regressieanalyse:

Vorige week ging het om 1 predictor variabele en 1 respons variabele. Dat ging over enkelvoudige regressieanalyse. Deze week gaat het om meerdere predictor variabele en 1 respons variabele. Dit gaat over multipale regressieanalyse. Dat betekent dat je met meerdere variabelen te maken hebt. Per individu zijn er hierdoor meerdere variabelen in je data te vinden.

De aanname bij deze analyse is dat binnen iedere subpopulatie de y-waarden normaal verdeeld zijn met gemiddelde  $\mu_y$  en standaarddeviatie  $\sigma$ .

De formule voor multipale regressieanalyse lijkt sterk op de formule voor enkelvoudige regressieanalyse. De formule is als volgt:  $\mu_y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_p X_p$

Waarom meervoudige regressieanalyse?

- De voorspelling is beter wanneer je meerdere variabelen gebruikt. Je hebt meer verklaarde variantie en minder residue variabelen.
- Voor als we casuale theorie willen onderzoeken en daarbij mogelijke superieure relaties willen elimineren.

### Multipale Regressie Analyse: grafisch

Grafisch gaat het nu niet om een regressielijn, maar om een regressievlak. De verticale afwijkingen in dit model zijn de residuen. Je hebt niet 1 regressielijn, maar bij iedere variabele hoort een andere lijn en daarom heb je een heel vlak nodig om grafisch de formule te kunnen representeren. Soms weet je al wat de alternatieve verklaringen zijn, maar soms moet je die uit je data halen.

### Correlatie bij multipale regressieanalyse:

Ook bij multipale regressieanalyse kun je een correlatie berekenen. Deze correlatie is complexer dan de enkelvoudige regressieanalyse. Er zijn namelijk meer variabelen bij in het spel.

Vaak valt  $R^2$  te hoog uit voor de populatie. SPSS berekent om deze reden altijd  $R^2$  *adjusted*. Dit is een effectmaat bij multiële regressie die bijgesteld is. Het zorgt ervoor dat  $R^2$  kleiner en realistischer uitvalt dan het normaal zou doen.

### **Het aantal vrijheidsgraden:**

Vorige week viel het aantal vrijheidsgraden al op:  $df$  was hier  $n-2$ . Nu kunnen we ook vertellen waarom dit zo is. Officieel is het aantal vrijheidsgraden bij zowel enkelvoudige als multiële regressieanalyse:  $df = n-p-1$ .  $P$  is in deze formule het aantal onafhankelijke variabelen.

Als we dit toepassen bij enkelvoudige regressieanalyse, waar men met 1 onafhankelijke variabele te maken heeft, krijg je de volgende formule:  $df = n-1-1 = n-2$ .

Bij multiële regressieanalyse zijn er meerdere onafhankelijke variabelen. Het verschilt per toets hoeveel variabelen er zijn en daarom is het aantal vrijheidsgraden:  $n-p-1$ .

Hoe meer predictoren, hoe minder vrijheidsgraden. Wanneer je een predictor toevoegt of verwijdert, verandert er van alles.

### **De ANOVA- tabel voor Multiële Regressieanalyse:**

Deze tabel is hetzelfde als de tabel voor enkelvoudige regressieanalyse. Er is 1 verschil: het aantal vrijheidsgraden. Dit is bij multiële regressieanalyse namelijk afhankelijk van het aantal onafhankelijke variabelen.

### **Het stelsel van hypothesen:**

Wanneer men met multiële regressieanalyse aan de slag gaat, moet men eerst een nulhypothese en een alternatieve hypothese opstellen. Voor deze vorm van regressieanalyse wordt het volgende stelsel van hypothesen opgesteld:

$$H_0: \beta_1 = \beta_2 = \beta_3 = \dots = \beta_p = 0$$

$$H_a: \text{Minstens 1 van } \beta_j \neq 0$$

### Voorbeeld uit het college:

Om de bovengenoemde stof toe te lichten staat hieronder een samenvatting van een voorbeeld dat werd gebruikt in het college. In dit voorbeeld wordt het vetpercentage voorspeld.

Het eerste wat gedaan moet worden is het opstellen van het stelsel van hypothesen:

- $H_0: VETP = BMI = LFT = SEXE = 0$
- $H_a: \text{Minstens 1 van de bovenstaande variabelen} \neq 0$

Met de data uit de SPSS uitvoer is de formule samen te stellen:

$$VETP = 1.11 BMI + 0.24 LFT - 9.17 SEXE - 3.46$$

Dan kunnen we het aantal vrijheidsgraden vaststellen. Voor multiële regressieanalyse was de volgende formule van toepassing:  $df = n - p - 1$ .  $Df = 286 - 3 - 1 = 282$ . Je kunt dit aflezen uit de SPSS uitvoer op sheet 20.

In de SPSS uitvoer is ook de F- waarde te vinden. Deze is te toetsen met het aantal vrijheidsgraden en blijkt erg significant.

Je conclusie is dus:  $H_0$  valt te verwerpen: minstens 1 van de variabelen is niet gelijk aan 0.

Subpopulaties: Mensen scoren dezelfde scores. Bijvoorbeeld in variabele 1: 3, variabele 2: 4 en variabele 3: 2. Een ander persoon scoort dezelfde scores, maar dan in een andere volgorde.

De assumpties bij regressie-analyse zijn:

- Alle variabelen moeten kwantitatief zijn.
- Gemiddelde respons heeft een lineaire functie
- De residuen zijn onafhankelijk
- De residuen zijn normaal verdeeld
- Homoscedasticiteit: Spreiding is hetzelfde. We willen dat de spreiding gelijk is voor elke subpopulatie. Boven de 3 of onder de -3 scoren minder dan we verwachten. Dit zijn de outliers.
- Heteroscedasticiteit: Spreiding is niet normaal.

## Week 3: Werkgroep

### Opdracht 3.1:

- Ystress en Xstab zijn tweezijdig significant. Voor een eenzijdige toetsing zijn ook alleen Ystress en Xstab significant (met een alpha van 0,05).
- $R^2_{x1,y} = 0,8856^2 = 0,784$ . Dit betekent dat 78,4% van de variantie wordt verklaard door X1.  $R^2_{x2,y} = 0,7271^2 = 0,5287$ , dus 52,87% van de variantie wordt verklaard door X2.
- Hij kan het beste beide predictorvariabelen gebruiken, want ze correleren laag met elkaar. Als de twee variabelen hoog gecorreleerd zouden zijn, geven ze dezelfde info en kan je ze beter niet beide gebruiken.

### Opdracht 3.2:

- De regressievergelijking met twee predictoren voor streekproefgegevens ziet er in het algemeen zo uit:  $\hat{y} = b_0 + b_1x_1 + b_2x_2$ .
- We gebruiken de ongestandaardiseerde regressiegewichten, dit geeft  $\hat{y} = 23,314 + 0,363x_1 + 0,059x_2$ .
- We gebruiken een t-toets met  $t = b_1 / SE_{b_1} = 0,363 / 0,049 = 7,38$ . Met een df van  $n - p - 1 = 2$  geeft dit een p-waarde tussen de 0,01 en 0,005. Ho moet dus worden verworpen. In de tabel in het werkboek staat dat de p-waarde 0,0179 is, maar dit komt omdat het een tweezijdige toets is (SPSS geeft de p-waarden altijd tweezijdig). Voor tweezijdig is het  $0,0179/2 = 0,00895$ . Dit ligt tussen de 0,01 en 0,005 dus het klopt.
- B2 is significant voor een eenzijdige toetsing. Er geldt dan namelijk  $0,0517 / 2 = 0,02585$ .
- Beide predictoren zullen een rol spelen in de populatie want ze zijn beide significant.
- Proefpersoon 2 heeft de waarden  $x_1 = 45$  en  $x_2 = 42$ . Dit geeft  $\hat{y} = 23,314 + 0,363 \cdot 45 + 0,059 \cdot 42 = 42,13$ .
- $(y - \hat{y})$  is het residu, de error. Als het residu negatief is, betekent dit dat er een onderschatting van de waarde is geweest. (als het residu positief is, is er dus een overschatting geweest).
- Een lijn zegt iets over één predictor, maar we hebben nu twee predictoren dus kan er geen lijn gemaakt worden.
- In plaats van een regressielijn kan er een vlak getekend worden. Het wordt dan een complex driedimensioneel figuur. SPSS kan dit niet uitvoeren, maar R kan dit wel. Met drie predictorvariabelen is het erg moeilijk maar het kan wel, je krijg dan een soort filmpje



Opdracht 3.3:

- a. Bereken via de tabel:  $SSE = \sum(y - \hat{y})^2 = 1,96$ .
- b. Gebruik de rekentip uit college twee.  $SST = \sum(y_i - \bar{y})^2 = (n-1) * s_y^2 = 4 * 22,50 = 90$ .
- c.  $SSM = SST - SSE = 90 - 1,96 = 88,04$ .
- d.  $SSM$  uitrekenen kan ook door  $(\hat{y} - \bar{y})^2$ . Hier komt ook 88,04 uit.
- e.  $H_0: \beta_1 = 0$  en  $\beta_2 = 0$  (kan ook geschreven worden als  $\beta_1 = \beta_2 = 0$ ).  
 $H_a: \beta_1$  of  $\beta_2 \neq 0$ .
- f. De nulhypothese wordt getoetst met een F-toets. F is hier 45,03.
- g. Zoek in de F-tabel de F-waarde op met  $df=2$  en  $df=2$ . De p-waarden liggen tussen 0,025 en 0,01. Met een alpha van 5% moet  $H_0$  dus worden verworpen.
- h.  $R^2 = SSM/SST = 88,04/90,00 = 0,978$ . Dit betekent dat het een sterk effect is.
- i.  $SSM = R^2 * SST = 0,978 * 90 = 88,02$ .  
 $SSE = (1 - R^2) * SST = 0,022 * 90 = 1,98$ .  
Er komt bijna hetzelfde uit, er is 0,02 verschil.
- j. R is de correlatie tussen de geobserveerde y-waarden en de voorspelde  $\hat{y}$ -waarden.
- k. We moeten nog steeds beide variabelen gebruiken.

Opdracht 3.4:

- a. De uitkomsten uit de SPSS-tabel en uit de berekeningen komen overeen.
- b. Voor de steekproef gebruiken we  $R^2$ . R zegt namelijk niks over VAF dus die niet en...
- c. De adjusted R square is voor de populatie. Dit is een benaderingsmanier.

## Week 4: College: Experimenteel onderzoek en experimentele controle

### Causaliteit:

Oorzaak en gevolgrelaties zijn de sleutel tot het begrijpen van de werkelijkheid.

Er zijn 3 criteria voor causaliteit:

1. Er moet samenhang zijn.
2. De oorzaak zal in tijd moeten voorafgaan aan het gevolg. Je moet dus de richting weten van de samenhang.
3. Je moet alternatieve verklaringen kunnen uitsluiten.

### Een verborgen variabele:

Om aan het laatste criterium te voldoen moet je alternatieve verklaringen kunnen uitsluiten. Dit is vaak erg lastig, omdat er vaak een derde variabele in het spel is. Dit noem je een verborgen of 'lurking' variabele. Deze verborgen variabelen kunnen de interne variabiliteit aantasten.

Met het volgende voorbeeld wordt het effect van een verborgen variabele duidelijk: onderzoekers vonden dat er een verband was tussen het aantal ijsjes dat verkocht werd en het aantal mensen dat verdronk: hoe meer ijsjes er verkocht werden, hoe meer mensen er verdronken. Echter bleek bij dit onderzoek dat het aantal uren zonneshijns beide variabelen verklaarden. Het aantal uren zonneshijns is dan ook de verborgen variabele. Het verband op deze manier met de 3 variabelen noem je een spurieuze correlatie.

### Confounding correlatie:

Wanneer zowel x als y causaal samenhangen als z en y. Dus y hangt causaal samen met x en met z.

Voorbeeld: opleiding (x) veroorzaakt inkomen (y), echter speelt ook het netwerk dat een persoon heeft een rol (z). Dus: inkomen wordt zowel veroorzaakt door opleiding, als door het netwerk wat een persoon heeft.

### Experimenteel onderzoek:

Om causaliteit te bewijzen moet men een experimenteel onderzoek doen. Hierbij kan men de volgende handelingen uitvoeren om causaliteit te bewijzen:

- Manipulatie van de onafhankelijke variabelen  
Een voorbeeld hiervan is het aantal milligram koffie dat je aan iemand toedient variëren. Dus aan groep 1 geef je 0 mg, aan groep 2 geef je 50 mg en aan groep 3

geef je 100 mg.

- Random toewijzing van de personen

Hierdoor worden groepen gelijkwaardig gemaakt. Je kunt dit op verschillende manieren doen. Het is belangrijk voor de validiteit van het experiment dat je gelijkwaardige groepen maakt.

- Invloed van de verborgen variabelen onder controle brengen

Dit doe je door alle variabelen onder controle te houden en dan te kijken wat varieert. Je zorgt er dus voor dat alles hetzelfde is, behalve 1 kenmerk.

### **Random toewijzing:**

Zoals we hiervoor zagen is dit om de validiteit te bewaren en groepen gelijkwaardig te maken. Je kunt de random toewijzing op de volgende 3 manieren toepassen:

1. Simpele random toewijzing: hierbij ga je gewoon random steekproeven trekken. Dit is totaal willekeurig.

2. Matched random toewijzing: hierbij ga je mensen matchen op kenmerken. Je kunt bijvoorbeeld matchen op intelligentie. Dan ga je groepjes maken van mensen die in dezelfde intelligentieklasse zitten.

3. Blocking random toewijzing: hierbij sorteert je mensen op een eigenschap, bijvoorbeeld geslacht, en ga je die mensen daarna random aan groepen toewijzen. Als je dit dus met geslacht doet, zie je dat in beide groepen evenveel mannen als vrouwen zitten.

Daarnaast is er nog de keuze of je een proefpersoon meerdere malen gebruikt, of dat je verschillende proefpersonen neemt voor verschillende condities. Dus laat je dezelfde proefpersoon op verschillende dagen de verschillende hoeveelheden koffie innemen, of deel je de proefpersonen in op de hoeveelheid koffie? *'Between-subjects design'* is een design waarbij verschillende proefpersonen onder de condities worden verdeeld. *'Within-subjects design'* is een design waarbij iedere proefpersoon in alle condities meedoet.

### **Bedreigingen interne validiteit:**

Als je onderzoek doet zijn er een aantal zaken waar je rekening mee moet houden.

Hieronder staan er een aantal op een rijtje:

1. Selectie: als je random toewijzing toepast, kan het je alsnog voorkomen dat er een bepaalde selectie is in een groep en dat een groep helemaal niet random is.

2. Uitval: zeker bij longitudinaal onderzoek kan het zo zijn dat mensen geen interesse meer hebben in het onderzoek en ermee stoppen. Dit maakt je resultaten minder betrouwbaar.

3. Testing: dit gebeurt als je voormeting toepast om te kijken of je de juiste manipulaties toepast. Hierbij kunnen mensen anders reageren na de voormeting doordat ze al blootgesteld zijn aan de variabelen.

4. Sensitatie: als er een voormeting is geweest, kunnen proefpersonen extra gevoelig worden voor de behandeling die toegepast wordt.
5. History: er kan een effect optreden door gebeurtenissen buiten de proefpersonen. Bij Local History is er een gebeurtenis van buitenaf in één van de groepen.
6. Maturation: veranderingen binnen de proefpersonen kunnen invloed hebben.
7. Instrumentatie: problemen met meetinstrumenten kunnen de interne validiteit beïnvloeden.
8. Diverse Design Confounds: dit is bijvoorbeeld het geval bij een onderzoeker die verschillende benaderingen heeft bij verschillende proefpersonen.

### **Verwachtingen:**

Zowel een onderzoeker als een participant kunnen verwachtingen hebben over het onderzoek. Om dit niet van invloed te laten zijn op de resultaten is het goed om een dubbel-blind onderzoek te maken. Hierbij weet de onderzoeker niet in welke fase de participant zich bevindt en kan de participant zich niet laten beïnvloeden door de onderzoeker. Ook het placebo-effect kan validiteitsproblemen opleveren.

### **Het dilemma van de onderzoeker:**

Als een onderzoeker meer experimentele controle wil hebben, wordt de interne validiteit vergroot, maar hierdoor wordt automatisch de externe variabiliteit verkleind. De onderzoeker staat hiermee voor een dilemma: wil hij zijn resultaten kunnen generaliseren of juist goede instrumenten gebruiken? Hij moet hier een afweging in maken.

## Week 4: Werkgroep

### Opdracht 4.1:

- a. De kenmerken van een experiment zijn: manipulatie van onafhankelijke variabelen, (random) toewijzen van proefpersonen aan condities en experimentele controle van externe variabelen.
- b. Leary beschrijft het within-subjects design. Dit is niet random, omdat dezelfde mensen steeds opnieuw in alle condities worden gebuikt.
- c. Dit is niet experimenteel te onderzoeken. Dit omdat er geen causaal verband kan worden aangetoond, maar ook omdat persoonskenmerken, zoals sekse, niet gemanipuleerd kunnen worden.

### Opdracht 4.2:

- a. Voor voorbeeld 1 geldt dat confounding door opvoeding een logische verklaring zou kunnen zijn. Voorbeeld 2 heeft een common respons voor leervaardigheid. Voorbeeld 3 heeft causatie als verklaring en voorbeeld 4 heeft een common response voor de ervaring van de therapeut.
- b. Deze vraag heeft betrekking op interne validiteit. Zijn er geen alternatieve verklaringen voor positieve correlaties?
- c. Een verborgen variabele is een mogelijke alternatieve verklaring als de verborgen variabele samenhangt met zowel de afhankelijke als de onafhankelijke variabele.

### Opdracht 4.3:

- a. In de experimentele groep zitten proefpersoon 1, 2, 3, 6, 7 en 10. In de controle groep zitten dan dus proefpersoon 4, 5, 8, 9, 11 en 12.
- b. In de experimentele en in de controle groep zitten evenveel mannen als vrouwen, er is een eerlijke verdeling.
- c. Maak een schema om de leeftijden in de verschillende groepen te vergelijken. Je ziet dan dat de leeftijden ongeveer overeen komen.
- d. We hadden een onafhankelijke t-toets kunnen gebruiken om vraag b en c te beantwoorden.
- e.  $HG_0(E) = 5,8$  en  $HG_0(C) = 6,3$ . Ze komen dus niet overeen, maar ze schelen maar een halve punt.
- f.  $HG(V) = 6,97$  en  $HG(M) = 5,17$ . Er is bijna 2 punten verschil en dus zijn het 'gevaarlijke' variabelen. Eigenlijk zouden we een t-toets moeten doen om erachter te komen of het wel een significant verschil is
- g. Sekse is een gevaarlijke variabele, maar mannen en vrouwen zijn wel gelijk verdeeld over de groepen. Ook zijn er metingen vooraf gedaan en de variabele leeftijd is niet gevaarlijk. Kortom, het resultaat van de randomisatie is geslaagd.

- h. Bij deze indeling van de groepen zitten alle vrouwen in de ene groep, en alle mannen in de andere groep. Dit is erg gevaarlijk voor de resultaten.
- i. Dan is het proces niet meer random. Blocking kan wel

Opdracht 4.4:

Bij blocking worden de onafhankelijke variabelen bij elkaar in groepen gezet. Er zijn dan dus 3 evengrote steekproeven, eentje met een laag opleidingsniveau, eentje met een middelbaar opleidingsniveau en eentje met een hoog opleidingsniveau.

Opdracht 4.5:

- a. Hoe hoger het opleidingsniveau, hoe meer afkeer er is ten opzichte van geweld.
- b. Ze zijn niet vergelijkbaar wat betreft opleidingsniveau. In de experimentele groep zitten bijvoorbeeld 3 laag opgeleiden, en in de controle groep zit er maar 1.
- c. De groepen zijn veel beter vergelijkbaar als je goede matches maakt.
- d.  $HG(E) = 6,13$  en  $HG(C) = 6,00$ . Ze zijn nu dus veel beter vergelijkbaar wat betreft houding ten aanzien van geweld.
- e. Als je de 2 mensen met het hoogste IQ laat matchen, en daarna de 2 mensen met het een na hoogste IQ.
- f. De praktische beperking is dat er soms niet erg goede matches tussen zitten. Je hebt dan dus eigenlijk een grotere  $n$  nodig.

Opdracht 4.6:

- a. In de experimentele groep zitten proefpersonen 3, 5, 7, 8, 10 en 12. In de controle groep zitten proefpersonen 1, 2, 4, 6, 9, en 11.  $HG1(E) = 5,8$  en  $HG1(C) = 6,9$ . Het verschil in  $HG1$  is dus  $6,9 - 5,8 = 1,1$ . Dat betekent dan mensen die een oorlogsfilm hebben gekeken toleranter zijn geworden ten opzichte van geweld.
- b.  $HG0(E) = 5,2$  en  $HG0(C) = 7$ .
- c. Er is geen erg groot verschil tussen de voor en de nameting. Het lag dus niet aan de film maar aan de houding die men al had voor het kijken van de film.

Opdracht 4.7:

- a. Met repeated measurements weten we zeker dat alle groepen gelijk zijn wat betreft persoonsgebonden kenmerken. Er is dan geen selectie meer.
- b. Dit is niet erg goed uitvoerbaar. We krijgen dan te maken met bijvoorbeeld carry over effecten.

## Week 5: College: Experimentele en Quasi-Experimentele Proefopzetten

### Soorten (Experimenteel) Onderzoek

Er zijn verschillende soorten onderzoek:

- Experimenteel onderzoek: controle over de manipulatie van de onafhankelijke variabele. Ook is er tot op zekere hoogte experimentele controle. Het gaat hierbij om 2 of meer groepen en via randomisatie worden proefpersonen over de groepen ingedeeld. De groepen zijn dus gelijk.
- Quasi Experimenteel onderzoek: hiervoor geldt hetzelfde als voor een experimenteel onderzoek, behalve dat bij een quasi experimenteel onderzoek geen randomisatie plaatsvindt.
- Pre experimenteel onderzoek: dit is geen goede manier van onderzoek doen, want hierbij vindt geen randomisatie plaats en ook is er geen of gebrekkige controle.
- Single-Case onderzoek: dit zijn verhalende beschrijvingen van bijvoorbeeld een individu of een groep. De informatie kan uit veel verschillende bronnen komen, bijvoorbeeld uit een dagboek of uit een interview.

### Echte experimenten: Terminologie

Met betrekking tot het aantal *Onafhankelijke* variabelen:

Als we met 1 onafhankelijke variabele te maken hebben spreken we van een éénweg experiment. Als we meerdere onafhankelijke variabelen bezitten spreken we van een meerweg experiment.

De aanwezigheid van niet-experimentele onafhankelijke variabelen kunnen bestaan uit compleet gerandomiseerde versus blokontwerp. Het blokontwerp wordt in het boek van Leary ook wel *expericorr* genoemd. Het gaat hierbij om minimaal 1 oorzakelijke variabele en die manipuleert een andere variabele die niet te manipuleren valt (bijvoorbeeld geslacht). Er is hier dus een niet experimentele 'onafhankelijke' variabele aanwezig. Als een blokvariabele categorisch is, is er geen probleem. Als een blokvariabele numeriek is (denk bijvoorbeeld aan leeftijden), moet je de 'median split' methode (geef rangnummers en deel de groep in tweeën) of de extreme groepen methode (neem alleen de twee uiterste 25% van de rangnummers) gebruiken. Helaas gooi je met deze methodes de echte scores weg, en dat is zonde. Een alternatief voor het blokontwerp is daarom de covariantie-analyse.

## **Variantiebronnen**

Er zijn 3 variantiebronnen:

1. Hoofdeffecten; effect van een enkele factor, hierbij kun je dus niet zomaar van interactie spreken.
2. Interactie effecten; interactie tussen de variabelen, hierbij zie je dat 2 variabelen met elkaar interacteren.
3. Toevalsvariantie; dit is ook wel de errorvariantie.

## **Quasi-Experimentele Ontwerpen**

*Nonequivalent Groups Design* en *Pretest-Posttest Design* zijn quasi-experimentele designs.

Hoe deze (en andere ontwerpen) in elkaar zitten, wordt geïllustreerd op de powerpoint

Voor quasi-experimentele designs zijn Sensitizatie (het effect wat iemand oploopt bij het ondergaan van een pretest) en local history (gebeurtenis die in de ene groep subject wel voorkomt, maar bij de andere niet en dus de resultaten beïnvloed) bedreigers van de validiteit.

## **Pre-Experimentele Ontwerpen**

Hieronder vallen *One-shot Case Study*, *One Group Pretest-Posttest Design* en *Nonequivalent Groups Posttest-Only Design*. Deze ontwerpen zijn geen goede manier van onderzoek doen. Er zijn hierbij veel gevaren, zoals history, maturation, testing en sensitizatie.

## **Longitudinaal onderzoek (*Panel Designs*)**

Bij longitudinaal onderzoek, onderzoek je een groep over een langere tijd. Er zijn wat negatieve aspecten aan deze manier van onderzoek: uitval kan plaatsvinden, het is lastig mensen te vinden die hieraan mee willen doen en het kost veel tijd en geld. Bij longitudinaal onderzoek is er dus geen variabele die je verandert, de tijd zelf is de quasi-independent variabele.

Om deze redenen bestaan er wat alternatieven voor longitudinaal onderzoek, zoals cross-sectioneel onderzoek. Hierbij onderzoek je mensen van verschillende leeftijden op dezelfde tijd. Echter heeft dit nadelen voor de betrouwbaarheid van het onderzoek: historie (ieder individu heeft een andere historie) en ontwikkeling (je neemt geen individuele ontwikkeling waar) zijn aspecten die moeilijk te controleren zijn.



## **Single Case Research**

Het probleem met het bestuderen van groepsgemiddelden is dat de gemiddelde verandering van alles kan betekenen, namelijk iedereen verandert, een deel verandert of verschillende delen veranderen verschillend. Hiervoor heeft men een oplossing: Single Case Studie.

Hierbij is observatie een sleutelwoord. Het is een gedetailleerde beschrijving van een individu, groep of gebeurtenis. De bedreiging bij deze manier van onderzoek is de observation bias. Er is namelijk maar 1 iemand die waarneemt en die kan dus subjectief waarnemen. Ook kunnen er geen causale verbanden uit dit onderzoek worden verklaard.

## **Single Case Designs**

- 1) ABA (treatment afgewisseld met antitreatment)
- 2) Multiple-I (verschillende treatments afgewisseld)
- 3) Multiple Baseline (Meer dan 1 gedraging bestuderen; aantonen dat alleen het 'targetgedrag' wordt beïnvloed)

## Week 5: Werkgroep

Opdracht 5.1.abc:

A: one shot case study.

X ----- O1

B: nonequivalent-groups design pretest only

1 NR X ----- O1

2 NR O1

C. one Group pretest-posttest design

O0----- X ----- O1

D. pretest-posttest-controlgroup design

1 R O0----- X ----- O1

2 R O0----- O1

E. posttest-only design

1 R X ----- O1

2 R O1

F. Solom vier groepen design

1 R O0----- X ----- O1

2 R O0----- O1

3 R X ----- O1

4 R O1

G. Dit is geen experiment, er is dus alleen een observatie.

Opdracht 5.1.d:

	A	B	C	D	E	F
Testing			X	S		S
Sensitisatie			X	X		S
History			X	S		S
Maturation			X	S		S
Uitval			X	X		X
Selectie		X		R	R	R + S

In de tabel is te zien welke gevaren er optreden bij welke onderzoeksvoorstellen. X betekent dat er gevaar is. S betekent dat het statistisch controleerbaar is. R betekend dat er geen selectiegevaar is door randomisatie.

Bij A is er maar één groep en maar één meting. Er kan niks vergeleken worden en dus kunnen er ook geen gevaren optreden. Het is net zoiets als een auto die zo veilig is dat hij niet kan rijden omdat de wielen er af zijn gehaald. Dit design is zo slecht dat er geen van deze gevaren kunnen optreden.

Bij B is alleen selectie een gevaar, de andere gevaren kunnen niet optreden omdat er geen voormeting is.

Bij C is alleen selectie niet een gevaar, omdat er maar één groep is.

Bij D is er geen selectiegevaar door randomisatie. Verder zijn alleen sensitatie en uitval een gevaar, want testing, history en maturation kun je berekenen.

Bij E is er geen test vooraf en kan er dus ook geen gevaar zijn. Er is geen selectiegevaar door randomisatie.

Bij F kan je alles berekenen behalve uitval. F is dan ook verreweg het beste onderzoeksontwerp.

(G staat niet in de tabel omdat deze gevaren bij G niet van toepassing zijn, omdat het enkel een observatie is.)

#### Opdracht 5.2:

- Experimenter expectancy houdt in dat de onderzoeker verwachtingen heeft, waardoor de participant beïnvloed kan worden.  
Er zijn twee soorten demand characteristics. 'Good subject' houdt in dat de participanten goed willen zijn, en dus extra hun best gaan doen voor wat ze denken dat de onderzoekers willen. Maar het kan zijn dat de proefpersoon een verkeerd idee heeft, of dat de onderzoekers juist de normale reacties willen zien. Het tweede is 'evaluation apprehension'. Hierbij krijgt een proefpersoon plankenkoorts omdat hij of zij zich bekeken voelt. Hierdoor presteren ze minder goed dan normaal het geval zou zijn.  
Het placebo-effect treedt op als iemand iets krijgt toegedient, bijvoorbeeld een pilletje, en dat diegene verwacht dat dit gaat werken. Hierdoor kunnen er daadwerkelijk dingen veranderen, ookal heeft het pilletje daar niks meer te maken. (Voor een placebo-effect is dus wel eerst een stimulus nodig.)
- Al deze bedreigers kunnen van invloed zijn op dit experiment.
- Een oplossing kan bijvoorbeeld een dubbel-blind onderzoek zijn. Ook kunnen ze proberen om het onderzoek breder te laten lijken, zodat de participanten niet weten dat het om nicotine gaat.

#### Opdracht 5.3:

- $O1(1) - O0(1) = 42 - 11 = 31$   $= X + HM + T + S$
- $O1(2) - O0(2) = 19 - 10 = 9$   $= HM + T$
- $A - b = 31 - 9 = 22$   $= X + S$
- $O1(3) - O1(4) = 36 - 16 = 20$   $= X$
- $C - d = 22 - 20 = 2$   $= S$
- De gemiddelde score op de voormeting gebruiken we als schatting voor de voormeting van groep 3 en 4. Deze wordt dan 10,5.  
 $O1(4) - O0(4)_{\text{geschat}} = 16 - 10,5 = 5,5$   $= HM$
- $B - f = 9 - 5,5 = 3,5$   $= T$
- Selectie is bij het solomon ontwerp geen bedreiger van de validiteit, omdat selectie door randomisatie en statistische controle onder controle is.

#### Opdracht 5.4:

- De grafiek hoort 4 lijnen te hebben (uitleg zie c).
- De bijspijker cursus werkt, er is een snelle stijging tussen week 4 en week 5.
- Nee, de groepen vertonen niet hetzelfde patroon. Groep 1 ligt ongeveer rond het gemiddelde. Groep 2 blijft hetzelfde scoren, die waren dus vanaf het begin al goed. Zij hebben geen bijspijker cursus nodig. Groep 3 blijft permanent stijgen.
- Bij groep 1 is er tussen week 1 en week 2 mogelijk een testing effect. Ook bij groep 3 vindt er testing plaats.
- In groep 3 vindt maturation plaats.
- De belangrijkste bedreiger hier is selectie. Het is hier erg duidelijk dat er vanaf het begin al verschillen zijn. Er is hier geen random toewijzing van toepassing en dat kan de uitkomsten van het onderzoek sterk beïnvloeden. Een oplossing hiervoor is een controle groep laten meedoen, die geen bijspijker cursus krijgen.

- g. Het gemiddelde in dit onderzoek laat zien dat de bijspijkerkursus effect heeft, terwijl het maar bij één van de 3 groepen werkt. Hier zie je dus erg goed dat een gemiddelde niet altijd een goede afspiegeling is van de werkelijkheid.

## Week 6: College: Eenweg Variantie-Analyse (ANOVA)

### Variantieanalyse en regressieanalyse:

Er is een verschil tussen variantieanalyse en regressieanalyse. Variantieanalyse onderscheidt zich van regressieanalyse doordat regressieanalyse een nominale predictor variabele bevat.

### ANOVA:

ANOVA is ook wel variantieanalyse. Hierbij worden gemiddeldes met elkaar vergeleken. De verschillen *tussen* de groepen (between-groups) worden vergeleken met de verschillen *binnen* de groepen (within-groups). Waar het bij ANOVA om gaat is het volgende:

DATA (totaal) = FIT (between) + RESIDU (within)

$x_{ij} = \mu_i + \alpha_i + \epsilon_{ij}$

$\alpha_i$  is een effectparameter, en die geeft aan hoeveel het groepsgemiddelde afwijkt van het groot gemiddelde.  $\alpha_i = \mu_i - \mu$ , dus som  $\alpha_i n_i = 0$

$X^*$  is: alle mensen in een subpopulatie hebben zelfde waarde.

$X$  is op intervalniveau, dus we kunnen er mee rekenen.

ANOVA wordt voorspeld via dummy variabelen. Hierbij gaat het dus maar om 2 groepen: een experimentele groep en een controlegroep.

Je hebt zowel eenweg- als tweewegtoetsen.

Bij éénwegtoetsen is er één onderscheid in groepen. Er is maar 1 predictor en dit lijkt veel op enkelvoudige lineaire regressie.

Bij meerwegtoetsen vergelijken we met méér predictoren. Hierbij zijn meerdere predictoren en dit lijkt veel op meervoudige lineaire regressie.

### Wanneer gebruiken we ANOVA?

Ten eerste is ANOVA te gebruiken bij een echt experiment. Bij een echt experiment kijk je of de behandeling effect heeft gehad.

Ook bij quasi-experimenten, observatie, studies en surveys is de ANOVA te gebruiken.

Hierbij kijk je of er verschil is binnen groepen.

### Gemiddelden Vergelijken

We kunnen niet zomaar het gemiddelde vergelijken. Dat de gemiddelden tussen verschillende condities verschillen is erg aannemelijk, maar het gaat er om of er een significant verschil is. Het zou namelijk gek zijn als 2 condities precies dezelfde gemiddelden hebben.

### **De F –toets:**

Omdat we weer met ANOVA te maken hebben, is de F-waarde gelijk aan  $t^2$

Op sheet 13 zijn de formules te vinden. F = de variatie tussen groepen gedeeld door de variatie binnen groepen.

### **Het Eénweg ANOVA model**

Het voorspellen van een score door groepsgemiddelde gaat wel voor éénweg maar niet voor meerweg.

De aanname bij dit model is dat de residuen normaal verdeeld zijn, dit lijkt ook weer op regressie. Ook moeten de residuen onafhankelijk zijn. Een goed opgezet design zorgt daarvoor.

In de praktijk heeft een experiment verschillende standaarddeviaties, maar wij gaan er nu vanuit dat de standaarddeviaties gelijk zijn. Hiervoor hebben we de volgende vuistregel: de grootste standaarddeviatie moet kleiner zijn dan 2 keer de kleinste standaarddeviatie. Dan kun je stellen dat de standaarddeviaties gelijk zijn.

### **De F-toets in ANOVA**

De nulhypothese die we gebruiken om de F-toets uit te voeren is heel restrictief, die zegt dat alle gemiddelden gelijk zijn. In de alternatieve hypothese stellen we dat er een verschil is in gemiddelden, er zijn dan tenminste 2 gemiddelden verschillend.

### **Procedures voor meervoudige vergelijkingen:**

Voor meervoudige vergelijkingen heeft men verschillende methodes. Wij gebruiken de Bonferroni methode. Hierbij ga je  $\alpha$  afstemmen op de meerdere toetsen. Je dit doet door  $\alpha/k$ . k is het aantal toetsen dat je gebruikt. Als je een  $\alpha$  van 5% hebt, deel je deze dus door k. Wanneer k =3, doe je  $0.05/3$  om op de bijgestelde  $\alpha$  te komen.



- o. Er kan hier geen correlatie gebruikt worden als effectmaat. Er zijn namelijk meer dan 2 condities.

## Opdracht 6.2:

a.  $x_{ij} = \mu + \alpha + \varepsilon$

b.  $\mu(\text{groot gemiddelde}) = 8 \cdot (17+20+28+35)/35 = 25$

$\alpha_1 = 17-25 = -8$

$\alpha_2 = 20-25 = -5$

$\alpha_3 = 28-25 = +3$

$\alpha_4 = 35-25 = +10$

Ho:  $\mu_1 = \mu_2 = \mu_3 = \mu_4$ . Ha:  $\neg$ Ho

Hier geldt, hoe hoger de score, hoe slechter. Negatieve affectparameters zijn in dit geval dus goed.

De aanname van homogeniteit klopt hier.

c.  $SSG = \sum n_i \cdot \alpha_i^2 = 8 \cdot ((-8)^2 + (-5)^2 + 3^2 + 10^2) = 1584$

$MSE = s_p^2 = SSE/DFE = \sum (n_i - 1) \cdot s_i^2 / \sum (n_i - 1) = 7 \cdot (6,71 + 6,24 + 7,12 + 6,65) = 6,68$

d.  $SST = SSG + SSE = 1584 + 187,04 = 1771,04$

e.	Df	SS	MS	F
Groups	3	1584	528	79,04
error	28	187,04	6,68	
Total	31	1771,04		

- f. Als we F opzoeken in de tabel zien we dat de p-waarde kleiner is dan 0,001.  
g. Ho moet dus worden verworpen.

## Opdracht 6.3:

a. 6 t-toetsen. Want  $(4 \cdot 3)/2 = 6$

b. Type 1 fout =  $1 - (1 - 0,05)^6 = 0,2649$ . Er is dan dus 26% kans op een type 1 fout.

- c. De 26% moet 5% worden. Dit doen we door de bonferroni correctie toe te passen.  
 $0,05/6 = 0,0083$ . Dit moet alpha worden en dat kunnen we controleren door: type 1 fout is nu:  $1 - (1 - 0,0083)^6 = 0,0490$ . De kans op een type 1 fout ligt nu weer onder de 5% en dat is wat we willen.



- d. We hebben een alpha van 0,0083. maar dit is tweezijdig dus voor  $t^{**}$  kijken we bij 0,00415. Dit staat niet in de tabel dus kijken we voor  $t^{**}$  bij 0,0025 maar nu met een df van 28. (dit is conservatiever) Hier vinden we een  $t^{**}$  van 3,047.
- e.  $MSD = t^{**} * \sqrt{MSE * \sum(1/n_i)} = 3,047 * \sqrt{6,68 * \sqrt{(1/8+1/8)}} = 3,9375$ .  
Als je daarmee de significantie van de verschillen toetst zie je dat er 5 verschillen significant zijn en eentje niet.
- f.  $CI(95\%) = (\text{gemstraf} - \text{gemmix}100\%) \pm MSD$   
 $CI(95\%) = -8 \pm 3,9375$   
 $CI(95\%) = [-11,94 \pm -4,06]$

## Opdracht 6.4:

- a.  $R^2 = SSG/SST = 4151,4282/4653,2958 = 0,8921$ . Dit is een groot effect.
- b. Rpb bepalen is niet zinnig, want er zijn meer dan 2 condities.
- c.  $\omega^2 = SSG - DFG * MSE / SST + MSE = 4151,4282 - (3*1,4298)/4653,2958+1,4298 = 0,8910$
- d. Er is geen effect van de steekproefgrootte op de effectmaten, het is alleen wel een verklaring voor het verschil tussen  $R^2$  en  $\omega^2$ .

## Opdracht 6.5:

- a. De assumpties bij ANOVA zijn dat de varianties homogeen moeten zijn, en dat de residuen normaal verdeeld en onafhankelijk moeten zijn. De onafhankelijkheid van de residuen kan je niet checken, dat zit hem namelijk in de proefopzet. Verder lijken de varianties niet homogeen verdeeld te zijn, en als je een grafiek maakt kun je zien dat de residuen ook niet normaal verdeeld zijn. De gegevens voldoen dus niet aan de assumpties voor de ANOVA.
- b. Mogelijke oplossingen zijn om te kijken of er uitbijters zijn (bijvoorbeeld de score van 19 bij conditie A, zou dit niet een tyfout kunnen zijn?). Ook kunnen we de scores transformeren (bijvoorbeeld logaritmisch) of we kunnen non-parametrische toetsen uitvoeren.

## Opdracht 6.6:

- a. In dit experiment is sprake van 5 condities. Dit kunnen we zien door naar de eerste vrijheidsgraden te kijken. Dit zijn er 4 en het is  $n-1$ . dus dan is  $n = 4+1 = 5$ .
- b. Er zijn 25 proefpersonen betrokken bij dit experiment.

- c. De kritische F-waarde voor een alpha van 0,01 zoeken we op in de tabel, en die geeft 4,43.
- d. Cocaine heeft dus effect op de Franse taalvaardigheid.

## Week 7: College: Tweeweg Variantie Analyse

### Het contrast

Als je weet dat twee  $\mu$ 's verschillen, wil je weten waar dat verschil zit. Je kan dan achteraf bekijken welke condities verschillen met behulp van multipale vergelijkingen. Maar als je van te voren al een idee hebt over het verschil, en er een hypothese over kan opstellen, dan kan je contrast toetsen gebruiken. Deze toetsen worden ook wel *a priori contrasten* of *planned comparisons* genoemd. De specifieke hypothese kan je op twee manieren opstellen. De eerste manier is door de eerste (controle) groep te vergelijken met het gemiddelde van de tweede en derde groep, je maakt dan van de 2 gemanipuleerde groepen één groep: één populatie. We krijgen dan  $H_0: \mu_1 = 0,5(\mu_2 + \mu_3)$ .  $H_a$  is dan:  $\mu_1 > 0,5(\mu_2 + \mu_3)$ . De tweede manier is door de controle groep weg te laten en de twee gemanipuleerde groepen met elkaar te vergelijken. We krijgen dan  $H_0: \mu_2 = \mu_3$  en  $H_a: \mu_2 > \mu_3$ .

De eerste manier kan ook makkelijker worden opgeschreven met het contrast  $\psi$ . Deze  $\psi$  is een combinatie van populatiegemiddelden die onder  $H_0$  gelijk is aan 0. Je haalt dan alles naar een kant van de vergelijking, dus dan krijg je  $H_0: \mu_1 - 0,5(\mu_2 + \mu_3) = 0$ . Dan is  $\psi = \mu_1 - 0,5(\mu_2 + \mu_3)$ . De nulhypothese wordt dan  $H_0: \psi = 0$ .

### Contrast toetsen

De contrastcoëfficiënten ( $a_i$ 's) bij de  $\psi$  van hierboven: 1, -0.5 en -0.5. De som van de  $a_i$ 's is gelijk aan 0, dit is altijd zo. Handig om te weten is dat je contrastcoëfficiënten kunt vermenigvuldigen, zodat het allemaal hele getallen zijn. In dit geval zou je kunnen vermenigvuldigen met 2. Je krijgt dan 2, -1 en -1. De uitkomsten van de toets blijven dan wel gewoon hetzelfde. Om contrasten te toetsen wordt een speciale t-toets gebruikt, die je berekent met de volgende formule:

$t = c / SE_c = \sum a_i x\text{-gem}_i / (s_p \sqrt{(\sum a_i^2/n_i)})$  met  $df = DFE = N - I$  (het aantal proefpersonen – het aantal condities). In deze formule is  $c$  het steekproefcontrast. We komen aan de gepoolde standaarddeviatie met de formule van vorige week:  $s_p = \sqrt{MSE}$ . We kunnen met deze toets dus eenzijdig toetsen door de vooropgestelde hypothese, dit is beter dan een post-hoc toets. Ook is er bij deze toets meer power.

### Tweeweg ANOVA

Bij meerdere onafhankelijke variabelen kun je twee x een éénweg ANOVA doen, maar een tweeweg ANOVA heeft meer voordelen. Ten eerste zijn de gegevens beter generaliseerbaar, de verscholen variabelen zijn namelijk onder controle. Ook is het onderzoek een stuk efficiënter, je hebt bijvoorbeeld de helft minder proefpersonen nodig. Daarnaast is de power groter, omdat de error variantie wordt verkleind door de tweede factor in het model. Er is dus

een groter deel van de variantie verklaard en dus is de power groter. Als laatste, en misschien wel het belangrijkste, geldt dat je bij een tweeweg ANOVA naar een interactie tussen twee variabelen kunt kijken.

De assumpties bij een tweeweg ANOVA zijn hetzelfde als die bij de eenweg ANOVA.

Het tweeweg ANOVA model ziet er zo uit:

$$\text{DATA} = \text{FIT} + \text{RESIDUAL}$$

$$x_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + \epsilon_{ijk}$$

Hier is  $i$  een conditie van factor A,  $j$  een conditie van factor B en  $k$  een proefpersoon.

$\alpha_i$ -dakje wordt geschat met  $x_{i\text{gem}} - x_{\text{gem}}$  en  $\beta_j$ -dakje met  $x_{j\text{gem}} - x_{\text{gem}}$ . De interactie-effecten  $\alpha\beta_{ij}$  (met dakjes) wordt geschat met  $x_{ij\text{gem}} - (x_{i\text{gem}} + \alpha_i\text{-dakje} + \beta_j\text{-dakje})$ . In een gebalanceerd design zijn de sommen van deze parameters gelijk aan 0.

De ANOVA tabel ziet er zo uit:

Source	DF	Sum of Squares	Mean Square	F
A	$I - 1$	$\sum n_i \hat{\alpha}_i^2$	SSA / DFA	MSA / MSE
B	$J - 1$	$\sum n_j \hat{\beta}_j^2$	SSB / DFB	MSB / MSE
AB	$(I - 1)(J - 1)$	$\sum n_{ij} \hat{\alpha}\hat{\beta}_{ij}^2$	SSAB / DFAB	MSAB / MSE
Error	$N - I \times J$	SSE	SSE / DFE	
Total	$N - 1$	$SST = \sum (x_{ijk} - \bar{x})^2$		

In de SPSS uitvoer van een tweeweg ANOVA (Analyze > General Linear Model > Univariate) staat een *Total* en een *Corrected Total*. Wij gebruiken de Corrected Total omdat daar het intercept niet wordt meegerekend.

### Effectgrootte

De proportie verklaarde variantie voor het totaal is  $SSM/SST$ . Als je de proportie verklaarde variantie wilt weten van een effect, dan krijg je  $\eta^2 = SS_{\text{effect}} / SST$ . Als we alleen het deel verklaard willen zien wat ook daadwerkelijk invloed heeft (dus zonder de variabele die geen invloed heeft en zonder de interactie) dan gebruiken we  $\eta^2_{\text{partial}} = SS_{\text{effect}} / (SS_{\text{effect}} + SSE)$ . Er geldt hier, hoe meer factoren, hoe kleiner  $\eta^2$  voor een bepaald effect. Bij de eenweg ANOVA is  $\eta^2$  gelijk aan  $\eta^2_{\text{partial}}$ . Als er dan ook maar 2 condities zijn geldt zelfs dat deze ook nog gelijk zijn aan  $r_{pb}^2$ .

Voor een schatting van de effectgrootte in de populatie krijgen we  $\omega^2\text{-dakje} = (SS_{\text{effect}} - (DF_{\text{effect}} * MSE)) / (SST + MSE)$ .